

Data Paper

A 16S rRNA amplicon dataset of Las Tablas de Daimiel National Park, Spain

Alejandro Rodríguez-Gijón¹, Víctor López-Maroto¹, Mireia Mestre¹, Salvador Sánchez-Carrillo¹, Rafael Laso-Pérez¹

¹ Department of Biogeochemistry and Microbial Ecology, Museo Nacional de Ciencias Naturales (MNCN-CSIC), Madrid, Spain

Corresponding authors: Alejandro Rodríguez-Gijón (alejandro.rgijon@mncn.csic.es); Salvador Sánchez-Carrillo (sanchez.carrillo@mncn.csic.es);

Rafael Laso-Pérez (rafael.laso@mncn.csic.es)

Abstract

Wetland ecosystems are of particular environmental importance due to their role as biodiversity shelters. However, the intense anthropization of wetland landscapes have propelled a decline in water quality and quantity globally during the last few decades. Under this context, microbial communities offer an accurate approach to better understand the conservation state of these water bodies. With the goal of contributing to environmental monitoring of wetlands of global significance, we present here a newly-generated 16S rRNA amplicon dataset derived from Las Tablas de Daimiel National Park, a wetland system in Spain designated as a Biosphere Reserve by UNESCO and a Ramsar site of international importance. Samples were collected across four different stations in the region, covering both the water column and the sediments, and environmental metadata were measured. In total, our dataset compiles 1,766 archaeal ASVs and 19,855 bacterial ASVs covering 11 and 64 phylum-level groups, respectively. This dataset offers a valuable resource for future monitoring work at the Las Tablas de Daimiel National Park to improve its health and conservation.

Key words: Freshwater ecology, metabarcoding, microbial ecology, wetlands



Academic editor: Florian Leese

Received: 24 September 2025

Accepted: 10 April 2026

Published: 4 May 2026

Citation: Rodríguez-Gijón A, López-Maroto V, Mestre M, Sánchez-Carrillo S, Laso-Pérez R (2026) A 16S rRNA amplicon dataset of Las Tablas de Daimiel National Park, Spain. Metabarcoding and Metagenomics 10: e173083. <https://doi.org/10.3897/mbmg.10.173083>

Copyright: © Alejandro Rodríguez-Gijón et al. This is an open access article distributed under terms of the Creative Commons Attribution License ([Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) – CC BY 4.0).

Introduction

Wetlands are vast areas of permanent or semi-permanent flooded land and represent approximately 13.4% of the global land area (excluding Antarctica) (Lehner et al. 2025). These ecosystems host a high number of macroscopic eukaryotic species, for instance serving as transient habitats for long-distance migratory birds (Junk et al. 2006) or as permanent habitats for rare and endangered plant species (Richardson et al. 2015). Hence, wetlands play a critical role as biodiversity shelters for both fauna and flora (Keddy et al. 2009). However, recent estimations report a loss of over 50% of the world's original wetland area (Davidson 2014), mainly driven by the increasing urbanization of the landscape (McCauley et al. 2013), the extensive development of agriculture (Asselen et al. 2013), and the cumulative pollution of the water (Prokisch et al. 2009). The critical state of global wetlands driven by anthropization highlights

the need for a joint effort towards environmental monitoring to improve the conservation of these ecosystems in order to halt their degradation.

Las Tablas de Daimiel National Park (TDNP) stands out amongst global wetlands since it represents one of the last river floodplains in southern Europe. The TDNP is a freshwater wetland of approximately 20 km² located in the center-south of the Iberian Peninsula, formed in the floodplain of the Cigüela and Guadiana rivers as a result of their seasonal overflows. Until the mid-1980s, it also functioned as the natural groundwater discharge of the Western La Mancha Aquifer, which is currently overexploited due to intensive groundwater extraction for irrigated agriculture (Sánchez-Carrillo and Angeler 2010). The TDNP is at the core of the UNESCO designated La Mancha Húmeda Biosphere Reserve (<https://www.unesco.org/en/mab/mancha-humeda>), and it is also a Ramsar designated wetland site (<https://www.ramsar.org>). However, during the past several decades, the wetland has suffered from intense groundwater pumping (Alvarez-Cobelas et al. 2001), high organic inputs from untreated wastewaters, and agricultural runoffs (Sánchez-Carrillo and Álvarez-Cobelas 2001; Santofimia et al. 2021), which has led to a hydrologically-fluctuating hypertrophic ecosystem with high concentrations of toxic trace elements in the sediments (Jiménez-Ballesta et al. 2017). Overall, these events have increasingly threatened the ecological functionality and biodiversity in the TDNP, making it essential to assess and monitor the ecosystem health of its wetlands to evaluate their conservation.

Amid the variety of predictors of ecosystem health, microbial communities have been widely used due to their quick response to anthropogenic impact across worldwide freshwater bodies (Meziti et al. 2016; Luo et al. 2022; Wang et al. 2022). In the TDNP, a few studies have offered an estimation of the abundance of various prokaryotic groups via fluorescence microscopy (Ortega-Mayagoitia et al. 2002; Rodrigo et al. 2003). However, only two studies have reported the genetic diversity of Archaea and Bacteria present in this ecosystem via 16S rRNA gene amplicon sequencing (D'Auria et al. 2010; Santofimia et al. 2024). This relatively poor characterization of the microbial communities at the TDNP, together with the differentiated taxonomical composition between pelagic and benthic microbial communities (D'Auria et al. 2010), highlight the need of a more complete spatiotemporal characterization of the microbial communities present in the TDNP wetlands. Future studies combining newly generated and publicly available 16S amplicon data, will provide a more complete view of the quality of the TDNP's waters.

With the goal of providing new valuable data to long-term monitoring efforts, we here release a dataset of six pairs of 16S rRNA amplicon libraries obtained from the Las Tablas de Daimiel National Park. In short, we collected six different samples in four different stations and sequenced two amplicon libraries per sample using specific primers to both Archaea and Bacteria. Since our broad ambition is to provide a comprehensive exploration of the microbial communities that are present in this park, we provide, together with the sequencing dataset, an overview of the number and diversity of the Amplicon Sequence Variants (ASVs) specific to each sample, and the total number of ASVs shared between samples. Furthermore, we provide a general view of the taxonomy of the different ASVs detected in our study and their relative abundance is also provided, as we explore the taxonomic similarity between samples.

Methods

Sampling and DNA extraction

In total, six different samples were collected across four different stations at the Las Tablas de Daimiel National Park (**TDNP**) on the 30th of May, 2023 (Fig. 1): Entradilla (samples collected from the water column and sediment, referred as “E_F” and “E_S”, respectively), Casilla de Benito (samples collected from the water column and sediment, referred as “CB_F” and “CB_S”, respectively), Puente Navarro (sample collected from the water column, referred as “PN_F”) and El Tablazo (sample collected from the sediment, referred as “T_S”, which was completely dry). Water samples were collected in 1-L polyethylene bottles (previously washed with 4% HCl) from the surface water of the different ponds, while sediment samples were collected by hand using a core sampler then stored in 50 mL polyethylene bottles. Environmental metadata and geographic location of the six collected samples are included in Suppl. material 1: table S1.

In parallel to sample collection, environmental parameters were measured *in situ*. Measurements of temperature and dissolved oxygen were conducted using a YSI ProODO probe, while pH and conductivity were measured with a Crison MM40. Major cations (Ca^{2+} , Mg^{2+} , K^+ , Na^+) were quantified by ICP-OES (Perkin Elmer Optima 4300 DV), and Cl^- and SO_4^{2-} by ion chromatography (Dionex 120). SiO_2 , CO_3^{2-} , and HCO_3^- were determined colorimetrically following the APHA (2005) protocols (American Public Health Association (APHA) 2005). Nutrients (NH_4^+ , NO_2^- , NO_3^- , and SRP) were analyzed using a SEAL AutoAnalyzer 3, following the APHA protocols (American Public Health Association (APHA) 2005). Chlorophyll-a was not measured due to sample freezing. Dissolved and total fractions of organic and inorganic carbon (DOC, DIC, TOC, IC) and total organic nitrogen (TON) were analyzed with a Shimadzu TOC-VCSH. Sediment carbon and nitrogen were measured with a Perkin-Elmer CHNSO-2400 elemental analyzer. Total phosphorus, after acid digestion, was measured using a SEAL AutoAnalyzer 3. Microbial biomass carbon, nitrogen, and phosphorus in surface sediments (0–5 cm) were extracted via chloroform fumigation-incubation (Sánchez-Carrillo et al. 2013), and quantified using the same instruments after cell lysis and acid digestion. Measured environmental parameters can be found in Suppl. material 1: table S1.

All samples were transported to the laboratory on the same day (temperature < 4°). Upon arrival, water samples were filtered using 0.22 mm Sterivex filters (Millipore) until silting of the filters. All filters were then stored in a -20 °C freezer until DNA extraction, except for sample “T_S”; due to the low hydration of the sample, we filled up the interstitial space of the sediment sample with LifeGuard Soil Preservation (Qiagen) before freezing to prevent DNA degradation. DNA extraction of all samples was performed using a DNeasy PowerSoil kit (Qiagen) following manufacturer instructions. After extraction, DNA concentration was measured for the six different samples using a Qubit device (ThermoFisher Scientific) to ensure a minimum DNA concentration prior to sequencing.

Library preparation, PCR conditions and sequencing

All samples were sent to the FISABIO Sequencing and Bioinformatics Service in Valencia, Spain for library preparation, PCR amplification, and amplicon

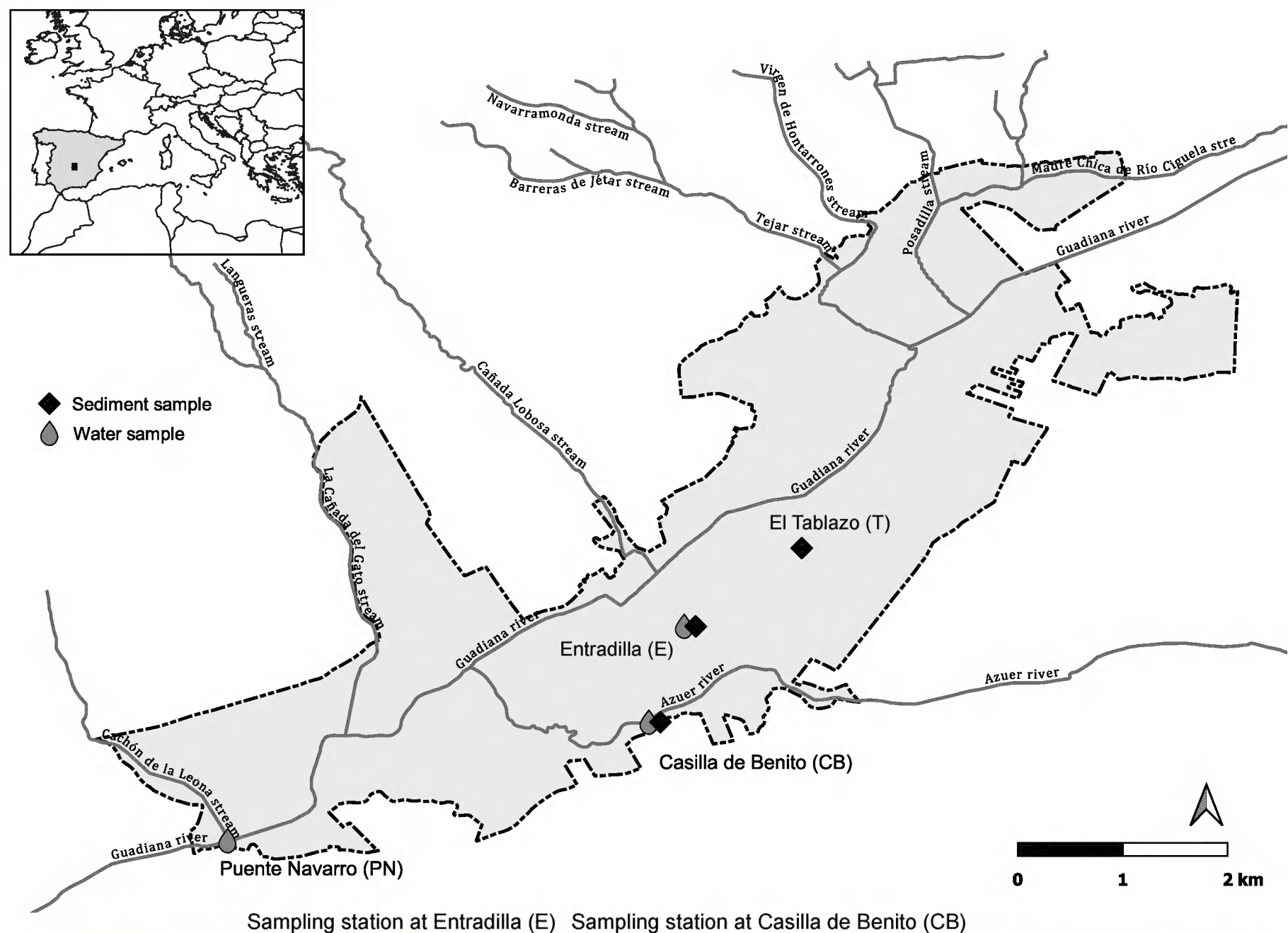


Figure 1. Location of sampling stations. The maps were created using QGIS v3.40.8 long-term release 'Bratislava'. The boundaries of the Las Tablas de Daimiel National Park were obtained from the geographic database of the Red de Parques Nacionales, provided by the Organismo Autónomo de Parques Nacionales of the Ministerio para la Transición Ecológica y el Reto Demográfico (last update: July 1, 2020). Surface watercourses and their toponyms were extracted from the metadata catalogue of the Infraestructura de Datos Espaciales (IDE) of the Hydrographic Confederation of the Guadiana (last update: June 22, 2025). Pictures on the bottom of the figure illustrate the sampling sites of Entradilla (E) and Casilla de Benito (CB).

sequencing. In summary, two amplicon libraries per sample of the 16S rRNA gene were created via PCR amplification for both Bacteria and Archaea.

A first amplicon PCR was performed to amplify our DNA samples using the specific primers with Illumina overhang adapters attached, including negative controls. No technical replicates were included. On the one side, the bacterial

primer pair Bact341F (5'-CCTACGGGNGGCWGCAG-3') (Klindworth et al. 2013) and Bact785R (5'-GACTACHVGGGTATCTAATCC-3') (Klindworth et al. 2013) were used to amplify the bacterial V3-V4 region. On the other side, the archaeal primer pair Arch349F (5'-GYGCASCAGKCGMGAAG-3') (Takai and Horikoshi 2000) and Arch915R (5'-GTGCTCCCCCGCCAATTCCT-3') (Stahl and Amann 1991) were used to amplify the archaeal V3-V5 region. PCR conditions were as follows: initial denaturation at 95 °C for 3 minutes followed by 25 cycles of annealing for the bacterial samples and 28 cycles of annealing for the archaeal samples (95 °C for 30 seconds, 55 °C for 30 seconds, and 72 °C for 30 seconds), and a final extension at 72 °C for 5 minutes using a KAPA HiFi HotStart ReadyMix (reference KK2602). Then, a second index PCR was performed to attach Illumina sequencing adapters and dual indices using the Nextera XT Index Kit (v2 FC-131-2001). PCR conditions were as follows: initial denaturation at 95 °C for 3 minutes followed by 8 cycles of annealing (95 °C for 30 seconds, 55 °C for 30 seconds, and 72 °C for 30 seconds), and a final extension at 72 °C for 5 minutes. Libraries were finally normalized to 4nM considering amplicon size and then pooled in the same volume of each library prior to sequencing.

The six pairs of amplicon libraries were sequenced using an Illumina Nextseq2000 P1 platform (reference 20075294). Sequencing was performed using an automated cluster generation and a paired-end sequencing (2×300 bp). Primer sequences were removed during the initial demultiplexing step performed by the sequencing facility and reads entering the DADA2 pipeline were already primer-free. The quality of the resulting raw reads was assessed using fastp v0.20.1 (Chen et al. 2018), setting a minimum read length of 50 bp, a mean minimum trimming quality of 20, and a sliding window of 10 bp.

Analysis of the 16S rRNA gene raw reads

After the delivery of the 16S rRNA gene sequences, they were analyzed in RStudio v2025.05.1+513 (R v4.5.1) (R Core Team 2020) using the DADA2 package v1.36.0 (Callahan et al. 2016). We first used the function “filterAndTrim” to test various trimming options for the raw reads. For the bacterial paired-end reads, we obtained the best preprocessing using “truncLen(270, 220), trimLeft=20”, while for the archaeal paired-end reads, it was “truncLen=c(290,290), trimLeft=17, maxEE=c(2,2)”. For further details about the different trimming methods assessed, please see: https://github.com/gecko1990/Daimiel_16S. The number of raw reads and trimmed reads per sample are included in Suppl. material 1: table S2.

In order to obtain an overview of the microbial community captured, we constructed an Amplicon Sequence Variants (ASVs) table from the trimmed reads using the “makeSequenceTable” function. The chimeric ASVs, including those classified as chloroplasts and mitochondria, were then eliminated using the “removeBimeraDenovo” function (method = “consensus”). To ensure that all chimeras were removed, we used the tool “—uchime_denovo” function from vsearch v2.17.1 to detect any remaining chimeras, but none were found (Rognes et al. 2016). To obtain an estimation of the biodiversity captured, we explored the number of ASVs detected according to our sequencing efforts by using the function “rarecurve” from the R package Vegan v2.7-1 (Oksanen et al. 2025). The diversity of ASVs shared between samples was explored by using intersection plots from the R package UpSetR v1.4.0 (Conway et al. 2017).

The taxonomic classification of the ASVs was performed using the DADA2 package v1.36.0 (Callahan et al. 2016). We leveraged the function “assignTaxonomy”, and used the SILVA rRNA database release v138.1 (Quast et al. 2012; Yilmaz et al. 2014) as reference. The relative abundance of each taxonomic group was then estimated using the R package phyloseq v1.16.2 (McMurdie and Holmes 2013).

To robustly explore ASV diversity and richness across samples, we performed a non-metric multidimensional scaling (NMDS) analysis using a Bray-Curtis distance, and visualized the results using ordination plots using the R package phyloseq v1.16.2 (McMurdie and Holmes 2013). In addition, we explored the present diversity using Shannon and Simpson diversity indices (Haegeman et al. 2013). These metrics were calculated using the “plot_richness” function from the phyloseq R package v1.16.2 (McMurdie and Holmes 2013).

Results and discussion

After sequencing, quality control, and removal of contamination and potential chimeras, our amplicon dataset totals 19,855 bacterial ASVs (grouped into 64 phyla; Suppl. material 1: table S3) and 1,766 archaeal ASVs (grouped into 11 phyla; Suppl. material 1: table S4). The mean amplicon length for archaeal ASVs is 386.04 bp (standard deviation = 81.43, range: 273–534 bp), while for bacterial ASVs it is 412.80 bp (standard deviation = 17.80, range: 250–428 bp). Interestingly, the prokaryotic diversity captured in our study is superior to previous amplicon surveys (D’Auria et al. 2010; Santofimia et al. 2024), stressing the scope of our analysis. More specifically, only 106 bacterial ASVs are common to all six samples, and they span over 12 different phyla; Proteobacteria (n = 31), Firmicutes (n = 29), Actinobacteriota (n = 13), Cyanobacteria (n = 11), Planctomycetota (n = 7), Bacteroidota (n = 4), Chloroflexi (n = 4), Coprothermobacterota (n = 2), Verrucomicrobiota (n = 2), Dependitiae (n = 1), Patescibacteria (n = 1), and Thermotoga (n = 1). For Archaea, only 2 ASVs classified to the phylum Crenarchaeota are common to all samples.

Despite the low number of ASVs shared between all samples, our amplicon dataset compiles a relatively large collection of ASVs across all stations. For instance, we detected the lowest number of ASVs in sample E_F (n = 2,000), where 15 archaeal ASVs and 299 bacterial ASVs are unique to this sample (Fig. 2). At the opposite end, the largest number of ASVs (n = 12,107) was retrieved from the sample CB_S, of which 709 archaeal ASVs and 3,485 bacterial ASVs are unique to this sample (Fig. 2).

Thanks to amplicon sequencing, microbiomes can be studied without the need for laboratory culturing. Our dataset indicates that up to 59.83% of the bacterial genera at the TDNP (i.e., 11,879 ASVs; Suppl. material 1: table S3) remain undescribed, while this proportion reaches up to 88.56% of the archaeal genera (i.e., 1,564 ASVs; Suppl. material 1: table S4). These values are consistent with previous surveys based on 16S rRNA gene amplicon sequencing, which have shown that the uncultured microbial majority can represent up to 81% of the microbial genera present across ecosystems (Lloyd et al. 2018). Furthermore, 421 bacterial ASVs and 255 archaeal ASVs could not even be classified at the phylum level, indicating that the TDNP might host a relatively large collection of unknown microbial taxa.

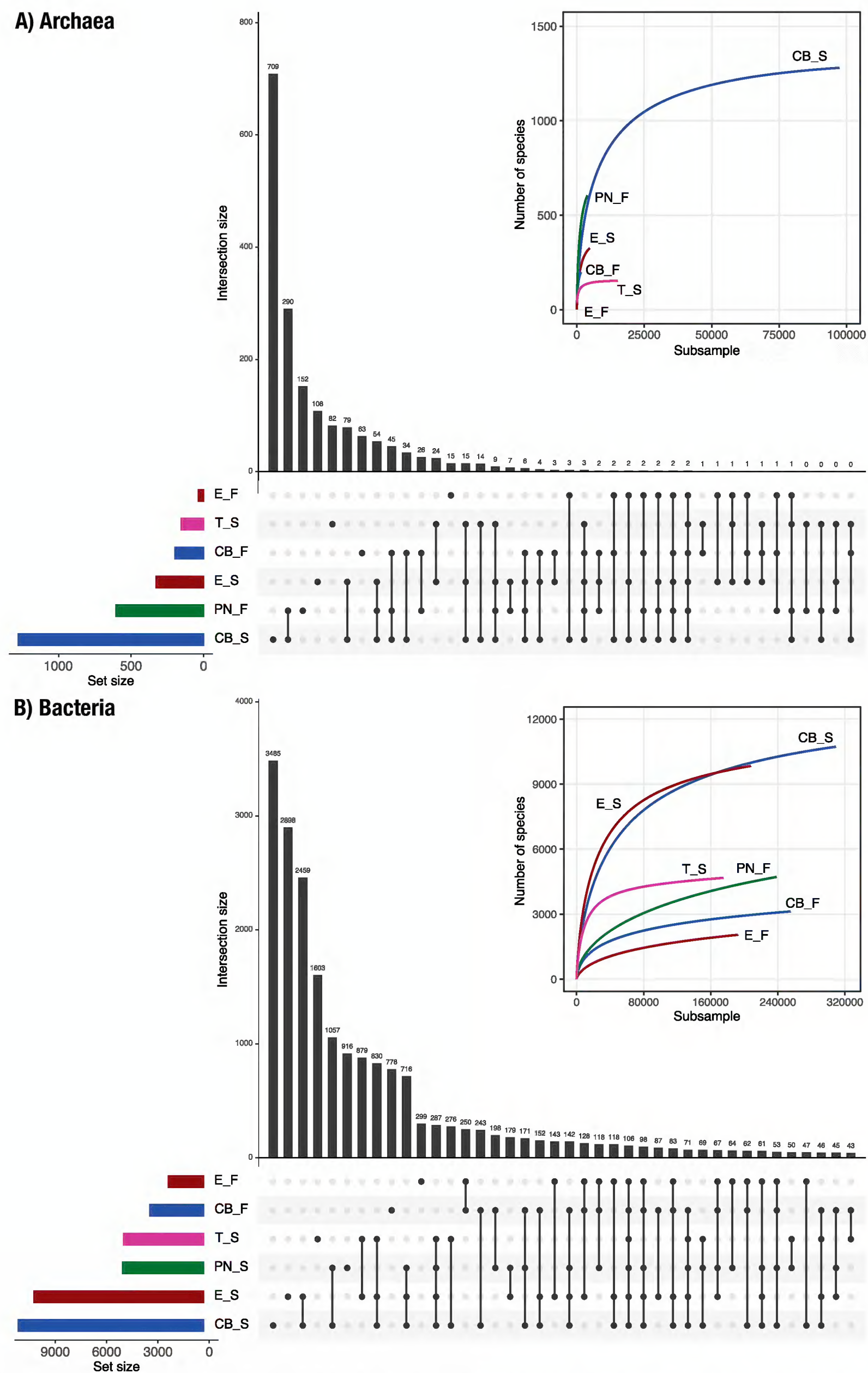


Figure 2. Upset barplot overviewing the number of Amplicon Sequence Variants (ASVs) shared between samples and unique to each sample. We include both archaeal ASVs (**A**) and bacterial ASVs (**B**). We designate samples in blue for the sampling station Casilla de Benito (CB_F and CB_S), in red for sampling station Entradilla (E_F and E_S), green for sampling station Puente Navarro (PN_F), and pink for sampling station El Tablazo (T_S).

In our amplicon survey, we also provide an overview of the relative abundance of the major prokaryotic class-level groups assigned to our ASVs per sample and domain. Overall, we observe a taxonomic diversification between pelagic and benthic microbial communities. For Archaea, the water column appears to be dominated by class Nanoarchaeia (mainly order Woesearchaeales) and various unclassified archaeal ASVs, while the most abundant class-level groups in the sediments are classified as Bathyarchaeia, Halobacteria, Nitrososphaeria, and Thermoplasmata (Fig. 3A and Suppl. material 1: table S4). Although differences in bacterial diversity are less noticeable, our results show that the water column is dominated by the classes Actinobacteria, Bacteroidia, Cyanobacteriia (order Synechococcales), and Gammaproteobacteria, while the most abundant classes in the sediments are Alphaproteobacteria, Cyanobacteriia (primarily orders Phormidesmiales and Cyanobacteriales), and other class-level groups with lower relative abundances (Fig. 3B and Suppl. material 1: table S3).

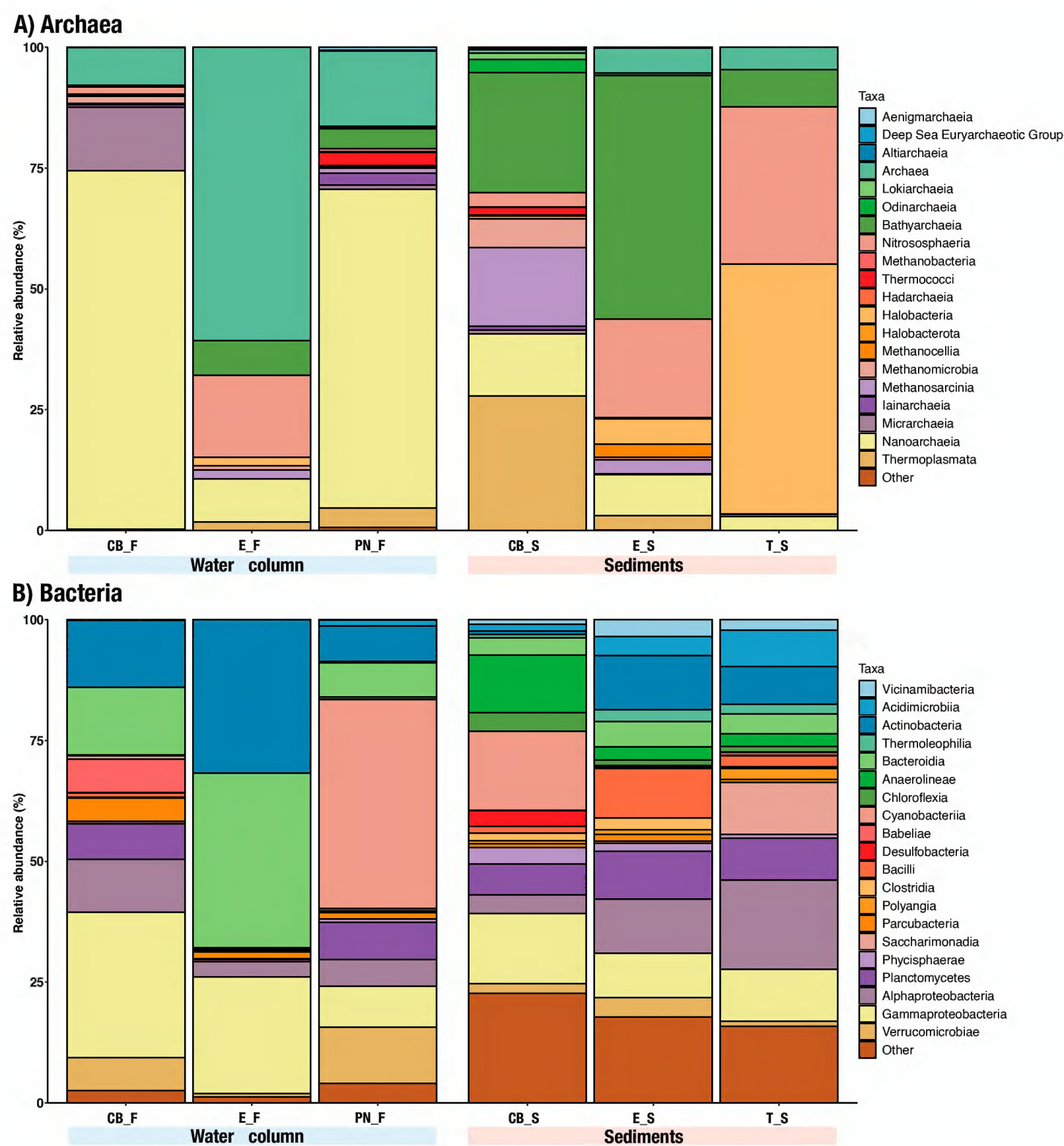


Figure 3. Taxonomic profiling of microbial communities per sample. We highlight the 20 class-level groups of highest relative abundance (%) for both Archaea (A) and Bacteria (B), and we incorporate the other classes with lower relative abundance in the category “Other”. Colors indicate the different taxonomic categories according to the legend on the right of the figure.

These differences in community composition between sediments and water column samples can also be observed when microbial richness is studied. We observe a clustering of benthic samples separately from pelagic samples for both Archaea and Bacteria according to NMDS metrics, while pelagic samples appear to be more dissimilar to each other (Fig. 4). Although the low stress values reported (stress = 8.3e-05 for Archaea and stress = 0 for Bacteria) could indicate a high fit of the data, these could potentially be influenced by the small sampling size of the study. To complementarily explore if pelagic and benthic communities differ, we estimated the Shannon and Simpson indices as robust metrics of taxonomic richness (Haegeman et al. 2013). These indices highlight a higher bacterial ASV richness in benthic microbial communities than in pelagic communities, while these differences are not as acute for archaeal richness (Fig. 4).

In conclusion, our 16S amplicon study complements previous amplicon surveys at the Las Tablas de Daimiel National Park by expanding the known benthic and pelagic diversity of Archaea and Bacteria. Furthermore, we observe

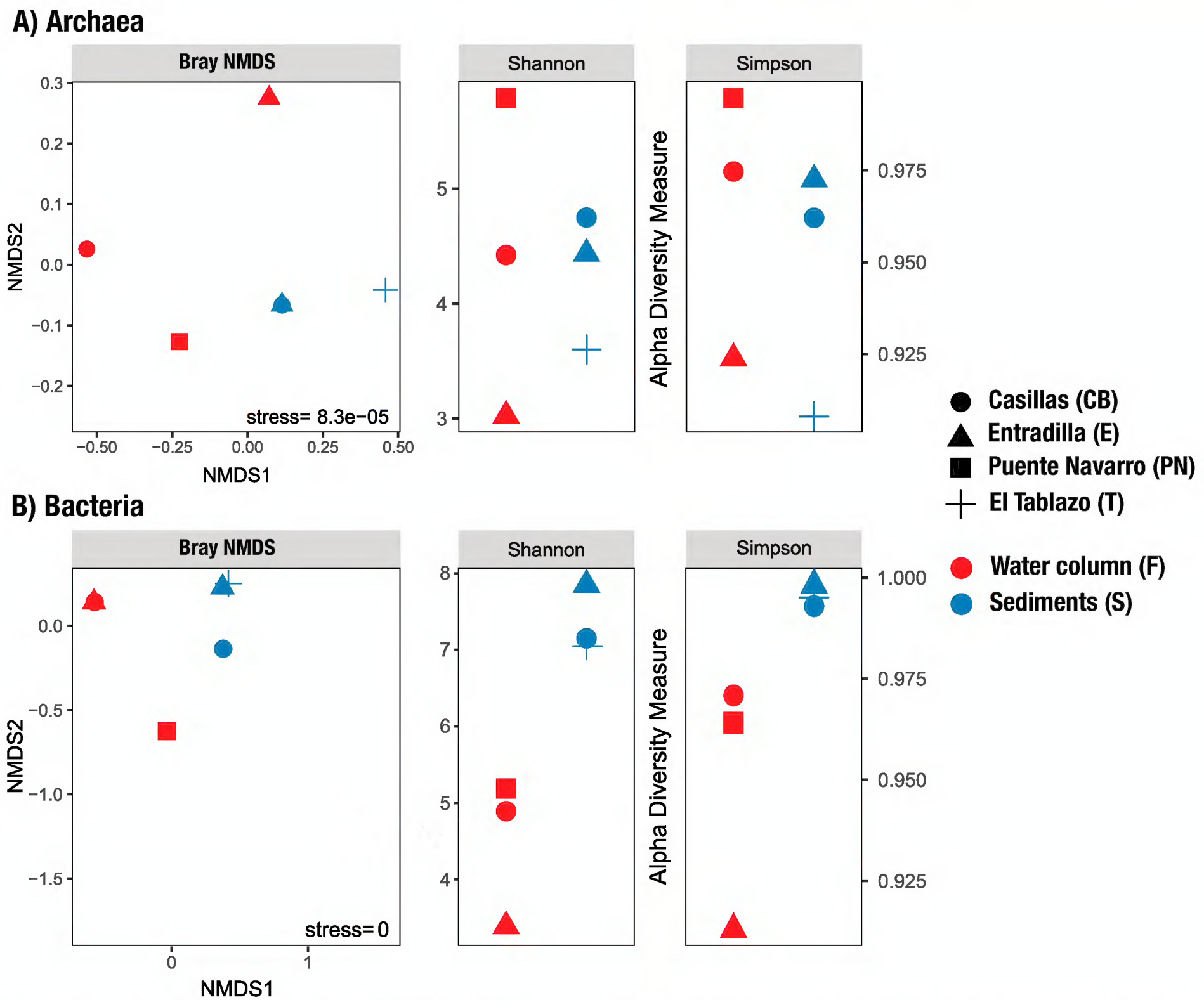


Figure 4. Overview of community composition and diversity in benthic and pelagic samples. We provide a summary of sample similarity using an NMDS analysis, and an estimation of the Shannon and Simpson indices to study prokaryotic richness for both Archaea (A) and Bacteria (B). Shapes indicate locations and colors type of sample according to the legend on the right of the figure.

an apparent taxonomic differentiation between benthic and pelagic samples. Lastly, the data here published offers a valuable resource for future monitoring efforts of the environmental health of this unique wetland ecosystem.

Technical validation

Library preparation, sequencing and the quality of the reads was monitored by the FISABIO Sequencing and Bioinformatics Service (Valencia, Spain). A detailed explanation of the different trimming approaches evaluated using the R package DADA2 (Callahan et al. 2016) to parse the 16S rRNA amplicon data can be found in https://github.com/gecko1990/Daimiel_16S.

Acknowledgements

We thank the staff of TDNP for their support and the facilities provided, and José Luis Ayala Liquiñano and Judit Osorio González, technicians at the Biogeochemistry Laboratory of the MNCN, for their assistance during field sampling and laboratory analyses.

References

- Alvarez-Cobelas M, Cirujano S, Sánchez-Carrillo S (2001) Hydrological and botanical man-made changes in the Spanish wetland of Las Tablas de Daimiel. *Biological Conservation* 97: 89–98. [https://doi.org/10.1016/S0006-3207\(00\)00102-6](https://doi.org/10.1016/S0006-3207(00)00102-6)
- American Public Health Association (APHA) (2005) 21 Standard Methods for the Examination of Water and Wastewater. American Public Health Association, Washington DC., 1288 pp.
- Asselen SV, Verburg PH, Vermaat JE, Janse JH (2013) Drivers of wetland conversion: A global meta-analysis. *PLoS ONE* 8: e81292. <https://doi.org/10.1371/journal.pone.0081292>
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* 13: 581–583. <https://doi.org/10.1038/nmeth.3869>
- Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34: i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Conway JR, Lex A, Gehlenborg N (2017) UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics* 33: 2938–2940. <https://doi.org/10.1093/bioinformatics/btx364>
- D'Auria G, Barón-Rodríguez M, Durbán-Vicente A, Moya A, Rojo C, Latorre A, Rodrigo M (2010) Unravelling the bacterial diversity found in the semi-arid Tablas de Daimiel National Park wetland (central Spain). *Aquatic Microbial Ecology* 59: 33–44. <https://doi.org/10.3354/ame01382>
- Davidson NC (2014) How much wetland has the world lost? Long-term and recent trends in global wetland area. *Marine and Freshwater Research* 65: e934. <https://doi.org/10.1071/MF14173>
- Haegeman B, Hamelin J, Moriarty J, Neal P, Dushoff J, Weitz JS (2013) Robust estimation of microbial diversity in theory and in practice. *The ISME Journal* 7: 1092–1101. <https://doi.org/10.1038/ismej.2013.10>

- Jiménez-Ballesta R, García-Navarro FJ, Bravo S, Amorós JA, Pérez-de-los-Reyes C, Mejías M (2017) Environmental assessment of potential toxic trace element contents in the inundated floodplain area of Tablas de Daimiel wetland (Spain). *Environmental Geochemistry and Health* 39: 1159–1177. <https://doi.org/10.1007/s10653-016-9884-3>
- Junk WJ, Brown M, Campbell IC, Finlayson M, Gopal B, Ramberg L, Warner BG (2006) The comparative biodiversity of seven globally important wetlands: A synthesis. *Aquatic Sciences* 68: 400–414. <https://doi.org/10.1007/s00027-006-0856-z>
- Keddy PA, Fraser LH, Solomeshch AI, Junk WJ, Campbell DR, Arroyo MTK, Alho CJR (2009) Wet and wonderful: The world’s largest wetlands are conservation priorities. *BioScience* 59: 39–51. <https://doi.org/10.1525/bio.2009.59.1.8>
- Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research* 41: e1–e1. <https://doi.org/10.1093/nar/gks808>
- Lehner B, Anand M, Fluet-Chouinard E, Tan F, Aires F, Allen GH, Bousquet P, Canadell JG, Davidson N, Ding M, Finlayson CM, Gumbrecht T, Hilarides L, Hugelius G, Jackson RB, Korver MC, Liu L, McIntyre PB, Nagy S, Olefeldt D, Pavelsky TM, Pekel J-F, Poulter B, Prigent C, Wang J, Worthington TA, Yamazaki D, Zhang X, Thieme M (2025) Mapping the world’s inland surface waters: An upgrade to the Global Lakes and Wetlands Database (GLWD v2). *Earth System Science Data* 17: 2277–2329. <https://doi.org/10.5194/essd-17-2277-2025>
- Lloyd KG, Steen AD, Ladau J, Yin J, Crosby L (2018) Phylogenetically novel uncultured microbial cells dominate earth microbiomes. *mSystems* 3: 10.1128/msystems.00055-18. <https://doi.org/10.1128/msystems.00055-18>
- Luo J, Zeng H, Zhou Q, Hu X, Qu Q, Ouyang S, Wang Y (2022) Anthropogenic impacts on the biodiversity and anti-interference ability of microbial communities in lakes. *Science of The Total Environment* 820: e153264. <https://doi.org/10.1016/j.scitotenv.2022.153264>
- McCauley LA, Jenkins DG, Quintana-Ascencio PF (2013) Isolated wetland loss and degradation over two decades in an increasingly Urbanized Landscape. *Wetlands* 33: 117–127. <https://doi.org/10.1007/s13157-012-0357-x>
- McMurdie PJ, Holmes S (2013) Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* 8: e61217. <https://doi.org/10.1371/journal.pone.0061217>
- Meziti A, Tsementzi D, Kormas Ar K, Karayanni H, Konstantinidis KT (2016) Anthropogenic effects on bacterial diversity and function along a river-to-estuary gradient in North-west Greece revealed by metagenomics: Diversity patterns along a river-to-estuary gradient. *Environmental Microbiology* 18: 4640–4652. <https://doi.org/10.1111/1462-2920.13303>
- NCBI BioProject (2025) NCBI BioProject. <https://www.ncbi.nlm.nih.gov/sra/PRJNA1291496>
- Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O’Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Borman T, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, Martino C, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J (2025) Vegan: Community Ecology Package. <https://vegandevs.github.io/vegan/>

- Ortega-Mayagoitia E, Rodrigo MA, Rojo C, Álvarez-Cobelas M (2002) Picoplankton dynamics in a hypertrophic semiarid wetland. *Wetlands* 22: 575–587. [https://doi.org/10.1672/0277-5212\(2002\)022\[0575:pdiahs\]2.0.co;2](https://doi.org/10.1672/0277-5212(2002)022[0575:pdiahs]2.0.co;2)
- Prokisch J, Széles É, Kovács B, Gyo’ri Z, Németh T, West L, Harper S, Adriano D (2009) Sampling strategies for testing and evaluation of soil contamination in riparian systems at the Tisza River Basin, Hungary. *Communications in Soil Science and Plant Analysis* 40: 391–406. <https://doi.org/10.1080/00103620802649039>
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41: D590–D596. <https://doi.org/10.1093/nar/gks1219>
- R Core Team (2020) R: A language and environment for statistical computing. <https://www.R-project.org/>
- Richardson SJ, Clayton R, Rance BD, Broadbent H, McGlone MS, Wilmshurst JM (2015) Small wetlands are critical for safeguarding rare and threatened plant species. *Applied Vegetation Science* 18: 230–241. <https://doi.org/10.1111/avsc.12144>
- Rodrigo MA, Rojo C, Álvarez Cobelas M (2003) Autotrophic and heterotrophic picoplankton in wetlands: Differences with Lake patterns. *International Review of Hydrobiology* 88: 464–481. <https://doi.org/10.1002/iroh.200310587>
- Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: A versatile open source tool for metagenomics. *PeerJ* 4: e2584. <https://doi.org/10.7717/peerj.2584>
- Sánchez-Carrillo S, Angeler DG [Eds] (2010) 2 Ecology of Threatened Semi-Arid Wetlands. Springer Netherlands, Dordrecht. <https://doi.org/10.1007/978-90-481-9181-9>
- Sánchez-Carrillo S, Álvarez-Cobelas M (2001) Nutrient dynamics and eutrophication patterns in a semi-arid wetland: The effects of fluctuating hydrology. *Water, Air, and Soil Pollution* 131: 97–118. <https://doi.org/10.1023/a:1011903300635>
- Sánchez-Carrillo S, Reddy KR, Inglett KS, Alvarez-Cobelas M, Sánchez-Andrés R (2013) Biogeochemical indicators of nutrient enrichments in wetlands: The microbial response as a sensitive indicator of wetland eutrophication. *Eutrophication: Causes, Consequences and Control*. Springer, Dordrecht, 203–222. https://doi.org/10.1007/978-94-007-7814-6_15
- Santofimia E, González-Toril E, De Diego G, Rincón-Tomás B, Aguilera Á (2024) Ecological degradation of a fragile semi-arid wetland and the implications in its microbial community: The case study of Las Tablas de Daimiel National Park (Spain). *Science of The Total Environment* 924: e171626. <https://doi.org/10.1016/j.scitotenv.2024.171626>
- Santofimia E, López-Pamo E, González-Toril E, Mejías M, Ruiz JM, Aguilera A (2021) Tablas de Daimiel National Park: The wetland versus the activity of invasive exotic species, the turbidity and eutrophication. *Geo-Temas*: 325–328.
- Stahl DA, Amann R (1991) Development and Application of Nucleic Acid Probes in Bacterial Systematics. *Nucleic Acid Techniques in Bacterial Systematics*. Wiley, Chichester, 205–248.
- Takai K, Horikoshi K (2000) Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes. *Applied and Environmental Microbiology* 66: 5066–5072. <https://doi.org/10.1128/AEM.66.11.5066-5072.2000>
- Wang Y, Guo M, Li X, Liu G, Hua Y, Zhao J, Huguet A, Li S (2022) Shifts in microbial communities in shallow lakes depending on trophic states: Feasibility as an evaluation index for eutrophication. *Ecological Indicators* 136: e108691. <https://doi.org/10.1016/j.ecolind.2022.108691>

- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, Da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 'T Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, Van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, Van Der Lei J, Van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR guiding principles for scientific data management and stewardship. *Scientific Data* 3: e160018. <https://doi.org/10.1038/sdata.2016.18>
- Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, Schweer T, Peplies J, Ludwig W, Glöckner FO (2014) The SILVA and “All-species Living Tree Project (LTP)” taxonomic frameworks. *Nucleic Acids Research* 42: D643–D648. <https://doi.org/10.1093/nar/gkt1209>

Supplementary material 1

Supplementary tables

Authors: Alejandro Rodríguez-Gijón, Víctor López-Maroto, Mireia Mestre, Salvador Sánchez-Carrillo, Rafael Laso-Pérez

Data type: xlsx

Explanation note: **table S1.** Sampling metadata of the samples collected. We include water body name, location and various environmental parameters measured. **table S2.** Overview of sequencing statistics provided by FISABIO of the six pairs of amplicon data. We include accession numbers. **table S3.** Overview of the bacterial ASVs retrieved, including their taxonomy and the number of reads classified as that ASV. **table S4.** Overview of the archaeal ASVs retrieved, including their taxonomy and the number of reads classified as that ASV.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: <https://doi.org/10.3897/mbmg.10.173083.suppl1>

Additional information

Conflict of interest

The authors have declared that no competing interests exist.

Ethical statement

No ethical statement was reported.

Artificial Intelligence (AI) use

The authors accept full responsibility for the content of the manuscript, including the disclosure of any use of AI.
No AI tools were used in the preparation of this manuscript.

Funding

This study was supported by the grant ‘Seguimiento Ambiental de Las Tablas de Daimiel – Limnología’ from the ‘Organismo Autónomo Parques Nacionales’ through TRAGSA number 674576. Additionally, it was partially supported by the Grant 2763/2021 funded by the Spanish Ministry of Ecological Transition and the Demographic Challenge through the Organismo Autónomo Parques Nacionales. RL-P and MM were funded by a Ramón y Cajal grant (RyC2021-031775-I for RL-P, and RyC RYC2021-031946-I for MM) from the Spanish Ministerio de Ciencia e Innovación (MCIN/AEI/10.13039/501100011033) and the European Union («NextGenerationEU»/PRTR). RL-P and AR-G were supported through the project ARGOS (Grant PID2023-147830NA-I00 by MCIN/AEI/ 10.13039/501100011033 and FEDER, UE). VL-M was funded by “Ayudante de Investigación” contract from the regional government of Madrid (PEJ-2024-AI/ECO-31856).

Author contributions

RL-P and SS-C conceived the project idea, and secured funding. RL-P and SS-C sampled the different water bodies at the TDNP, and RL-P extracted the DNA. RL-P and MM performed the computational analyses. AR-G and VL-M prepared the final figures for the manuscript. AR-G submitted all genomic data to public repositories and wrote the first draft of the manuscript. All the authors contributed, with significant inputs, to the final manuscript and have read and approved the final version of the manuscript.

Author ORCIDs

A. Rodríguez-Gijón  <https://orcid.org/0000-0002-1649-6894>
V. López-Maroto  <https://orcid.org/0000-0002-8590-0395>
M. Mestre  <https://orcid.org/0000-0003-0986-633X>
S. Sánchez-Carrillo  <https://orcid.org/0000-0002-5471-977X>
R. Laso-Pérez  <https://orcid.org/0000-0002-6912-7865>

Data availability

The publication of the genetic data we generated followed the FAIR Data Principles for scientific data management (Wilkinson et al. 2016). The raw data generated in this manuscript are available under BioProject accession number PRJNA1291496 (NCBI BioProject 2025) in NCBI, and their corresponding BioSample IDs can be found in Suppl. material 1: table S2. Environmental metadata is available in Suppl. material 1: table S1. All other data that support the findings of this data paper are found in the main text and supplementary materials.